

Hybrid Inverse Motion Control for Virtual Characters Interacting with Sound Synthesis

Application to Percussion Motion

A. Bouënard
S. Gibet
M. M. Wanderley

Received: date / Accepted: date

Abstract The ever growing use of virtual environments requires more and more engaging elements for enhancing user experiences. Specifically regarding sounding virtual environments, one promising option to achieve such realism and interactivity requirements is the use of virtual characters interacting with sounding objects. In this paper, we focus as a case study on virtual characters playing virtual music instruments. We address more specially the real-time motion control and interaction of virtual characters with their sounding environment for proposing engaging and compelling virtual music performances. Combining physics-based simulation with motion data is a recent approach to finely represent and modulate this motion-sound interaction, while keeping the realism and expressivity of the original captured motion. We propose a physically-enabled environment in which a virtual percussionist interacts with a physics-based sound synthesis algorithm. We introduce and extensively evaluate the Hybrid Inverse Motion Control (*HIMC*), a motion-driven hybrid control scheme dedicated to the synthesis of upper-body

percussion movements. We also propose a physics-based sound synthesis model with which the virtual character can interact. Finally, we present an architecture offering an effective way to manage heterogenous data (motion and sound parameters) and feedback (visual and sound) that influence the resulting virtual percussion performances.

Keywords Physics-based Computer Animation · Motion Control · Motion-Sound Interaction · Sound Synthesis

1 Introduction

Playing a musical instrument involves complex human behaviors. Instrumental gestures are indeed gestures that require extreme rapidity, accuracy, and physical engagement with the instrument. While performing, a skilled musician is able to accurately control his motion, as well as perceive both the reaction of the instrument to his own actions and the resulting sound. The tight relationship between sound and the performer actions is crucial for understanding the mechanisms underlying musical learning and performance, and for finally proposing realistic and engaging virtual music performances. Transposing these real-world experiences into virtual environments offers the possibility of exploring novel solutions where virtual characters interacting with virtual musical instruments are central. We believe that these simulation experiments provide valuable clues for better modeling of instrumental gestures, and by extension for designing multimedia environments where animated entities interact realistically with sounding objects.

Most studies that focus on animating expressive and dexterous gestures use data-driven kinematics approaches, but these fail in modeling collisions that are involved in movements. However, the fine control of gesture-sound interaction with relevant musical nuances requires the modeling of effects due to human motor control. To compensate the lack of physical realism of these pure kinematic approaches, a class of methods for animating virtual characters that interact realistically with the environment propose physical approaches. Such approaches rarely address the simulation of instrumental gestures, and colliding effects to the corresponding sound production. Finally, simulating virtual characters interacting with musical instruments entails combining the principles of both kinematics and dynamics approaches, as a means of reflecting and taking into account the richness of motion strategies during music performances. In this work, we more specifically address the influence of motion control on the produced

A. Bouënard
McGill University, CIRMMT-IDMIL Lab.
555 Sherbrooke West, H3A 1E3 Montreal, Qc., Canada
Phone: [+1] 514-954-7146
E-mail: abouenard@gmail.com

S. Gibet
University of Bretagne Sud, VALORIA-SEASIDE Lab.
Campus de Tohannic, 56000 Vannes, France

M. M. Wanderley
McGill University, CIRMMT-IDMIL Lab.
555 Sherbrooke West, H3A 1E3 Montreal, Qc., Canada

sound, and we highlight three major issues in the case of percussing gestures:

- i) we adopt a physics-based approach, both for animating the virtual character and for synthesizing sound; we also model the subtle physical interactions occurring as the drumstick makes contact with the drum membrane. We assume that this physical modeling gives rise to compelling contact information that is shown to be one of the sources of motion-sound realism and variations [17];
- ii) we use motion capture data for driving our simulations. This type of control ensures that the produced motion remains fluid and expressive, and also provides significant clues for constraining the produced movements;
- iii) at a planning level, we focus on the preparatory motion which is very relevant in percussion practice [8]. The short duration of beat impacts moreover emphasizes the existence of motion strategies to rapidly form adequate contact profiles for various musical nuances [36].

The major contributions of this paper for synthesizing compelling percussion motion are in the direction of these three points. We propose a novel motion control approach, the so-called Hybrid Inverse Motion Control (*HIMC*), which combines a physics-based controller and a kinematics planner driven by motion capture data. Although this motion control scheme uses existing components from the animation community, its originality comes from the embedded nature of kinematic and dynamic controllers. Furthermore, we use at the planning level a reduced set of three-dimensional captured data, thus enabling a more intuitive conducting of the virtual character. Our aim is to show the feasibility, from an animation point of view, of a hybrid motion control scheme that integrates the inherent constraints of human data. We also propose an original real-time framework for animating the virtual character, and interacting with a sound-producing system. This multisensory environment is used to explore the interaction between motion and sound, by modifying the parameters at different levels: planning, motion and sound control. This opens new perspectives for a compositional system of gestures, editing gesture trajectories and producing in real time visual, audio and gestural outputs.

The paper is organized as follows. We begin in section 2 by providing an overview of our multisensory system, and exposing our motivations, for motion control,

sound synthesis and multisensory framework. Section 3 describes our *HIMC* scheme, an evaluation is also conducted by comparing qualitative and quantitative results with traditional inverse kinematics and inverse dynamics strategies. Finally, in section 4, we propose an interaction scheme with sound synthesis, and present simulation results. A discussion is outlined in section 5.

2 Overview and motivation

Three main issues are generally addressed when considering the animation and interaction of virtual characters with sound synthesis processes. One first issue is to preserve the realism of the simulated motion in comparison with captured motion examples, while enabling the physical interaction of the virtual character with its surrounding environment. It generally involves physically-based simulations driven by motion capture data, to ensure the physical integrity of the simulated motion even under external disturbances. A second critical issue is the sound synthesis technique to be used, so that a compelling and effective interaction can be defined between motion parameter outputs and sound parameter inputs. Finally, a third issue to be considered is the synchronization of these different modalities – motion and sound – and its use in interactive applications. Finally, we conclude this section by highlighting the specificities of drumming gestures.

2.1 Motion control

The adaptive and responsive control of virtual characters has been intensively investigated in physics-based computer animation research. Most of contributions have addressed the control of articulated figures using robotics-inspired proportional-derivative (*PD*) controllers [29]. This has inspired many works for handling different types of motor tasks, such as walking, running, swimming or boxing [19, 38, 40]. A lot of contributions have then focused on the main three issues concerning physics-based animation using controllers: reusability, tuning, and low/high level control. The first two issues are generally addressed either by composing controllers dedicated to specific motor tasks [16, 12, 14], or helping to cope with the hard and time-consuming process of tuning such *PD* controllers [1].

Our work is more related to the third issue, *i.e.* how to provide higher-level control policies in physics-based animation. This approach is motivated by the fact that most physical controllers use multi-dimensional postures as input, which prevents from intuitive control.

One solution is the use of hybrid methods, combining physics-based controllers and motion data, in such a way that we may associate the advantage of user controllability of kinematic methods and the responsiveness of dynamic controllers. Previous contributions have followed this direction. A solution is to use procedural methods combined with physics simulation [22]. Transitioning between kinematic and dynamic controllers has also been studied, either by using simple sequencing policies, or by using searching/indexing methods [32,40] based on motion capture data. Another recent contribution is the use of approximated equations of motions [25].

Our work differs from previous works by proposing a higher-level and intuitive control scheme, composed of a cascaded collaboration between inverse kinematics (*IK*) and inverse dynamics (*ID*). More specifically, this new method provides a solution to the control of a physics-based virtual character only by specifying end-joint positions. *IK* has also been used in conjunction with *ID* control in [39], but only as a pre-process step rather than in the core of the control algorithm. Regarding percussion motion, the use of captured data of reduced dimension as inputs to our *HIMC* control scheme is particularly consistent, since:

- i) drumstick extremity trajectories have been identified as motion parameters that characterize motion variability and may be used to discriminate different performances [8];
- ii) cognitive experiments show that the shape of the three-dimensional mallet trajectory prefigures the type of stroke and the nature of the produced sound [13];
- iii) such higher-level representation of motion is closer to the planning level than the whole posture of motion, and facilitates the edition of motion, according to a motion score composed of beat sequences [7,9].

Our approach for controlling the virtual character uses motion capture data. Driving the simulation by real motion data is a means of preserving the variability and expressiveness of real percussion gestures. We have already shown that expressive cues may be retained by drumstick extremity trajectories [8] and are due to various percussion playing variations, such as using different drumstick grips, motion variations reflecting musical nuances, as well as various beat impact locations.

We propose in this paper an original control strategy extracting end-joint trajectories to drive a combined inverse kinematics and dynamics systems, as shown by

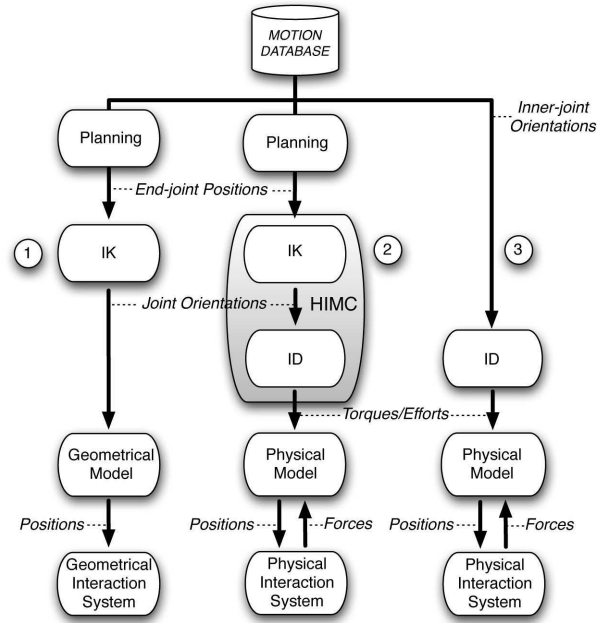


Fig. 1 Pipelines used for motion control: (1) Inverse Kinematics (*IK*); (2) Inverse Kinematics + Inverse Dynamics (*HIMC*); (3) Inverse Dynamics (*ID*). Commented dashed lines illustrate the data used for each pipeline.

pipeline 2 in figure 1, and partially introduced in [5]. Such a dual inversion process has not been addressed for motion control in computer animation. In previous contributions, traditional inverse kinematics applied to a geometrical model of a virtual character has been proposed, as illustrated by pipeline 1 in figure 1, but it does not allow physical interactions with a virtual instrument. Other approaches such as inverse dynamics directly exploit inner-joint orientations to drive a physical model, as shown by pipeline 3 in figure 1 [39, 4], but unfortunately offer a lower control level due to the high dimension of input motion poses. Conversely, our dual inversion scheme (*HIMC*) provides a higher-level control still with physical interaction. *HIMC* is particularly uneasy to model and implement, as it provides potential instabilities and tuning difficulties. In our approach, we avoid these instabilities by exploiting captured data to statistically determine the joint limits and to compute biomechanical parameters (center of mass, length, mass, density and inertia of the bodies). We also provide a comparative assessment of our *HIMC* model with conventional ones: classical inverse kinematics (*IK*), inverse dynamics (*ID*) from postures, and real human data.

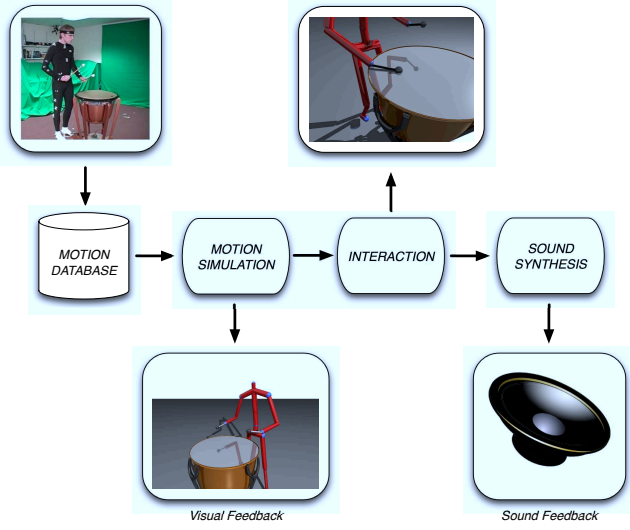


Fig. 2 System architecture and multimodal outputs. The *Motion Simulation* step involves a *Motion Capture Database* and results in the *Visual Feedback*. The *Interaction* expresses the mapping between the motion simulation and the *Sound Synthesis*, which results in the *Sound Feedback*.

2.2 Sound synthesis

Along with the dynamic animation of virtual characters, physically-motivated sound synthesis methods have been widely studied, with the analogous goal of creating sounds that can adapt to actions and changes in environment or object properties. Since the early interest of the graphics community in modeling the vibrations of objects for sound synthesis, most works have either focused on the direct use of vibrational modes based on measurements [34,11], or the simulation of surface vibrations using the finite element method [26]. More recently, the high interest on modal synthesis [23] has led to acceleration algorithms for handling complex environments or objects [28,3,27]. Our work involves in a similar manner physical models based on modal synthesis, but we point out and address the difficulty of relating these sound synthesis schemes to their cause in the context of music performance, *i.e.* the instrumental motion of a virtual performer. Here the focus is therefore on the simulation of percussion gestures, as well as on the interaction between motion and sound synthesis.

2.3 Multisensory architecture

We propose a real-time architecture for exploring the physical interaction between motion and sound synthesis. Figure 2 shows a block diagram of our framework. It is divided into three main parts: the motion simulation of the virtual character, the interaction process, and the sound synthesis process. Biomechanical data

are first extracted from the motion database and used to initialize the physical model of the virtual character. Motion capture data are also the input of the control policy involving our *HIMC* model. The motion simulation of the virtual character is then used to realize interacting experiments with a simulated sound synthesis system.

Our framework synchronizes different modalities and heterogeneous data: captured motion, motion simulation and sound control parameters used by the simulation processes; visual and audio data used to visualize the results of the simulation. Previous contributions have also adopted a motion-driven approach for the synchronized generation of soundtracks from animations [33,10]. However as demonstrated in many audio-visual works, despite many recent improvements, most computer animation and simulation frameworks are still reluctant to integrate such interaction or synchronization methods, therefore moving away from the interactive realism and presence that sound can give to computer animation applications. Our goal is to show how physical modeling of the motion-sound interaction, combined with a carefully-designed system architecture can be of interest to obtain engaging and realistic virtual environments. Moreover, we do not focus on a specific sound synthesis model, but we explore different interaction schemes attached to modal synthesis principles.

2.4 Specificities of percussion motion

There is a large amount of percussion instruments, which vary in size, shape, quality of the drumming surface. There is also a great variability among drumming gestures, such as the placement, the involvement of the whole body or not, the force needed for striking, and the type of the grip. Percussion performances involve ballistic gestures that precede the strike, the so-called "preparatory gesture", and impulse-based gestures during strike. Moreover, the contact with the instrument is limited to a few millisecond [36], consequently the striking force and effects on the produced sound have to be induced by the entire gesture. Finally the use of a striking tool such as a drumstick, a mallet, a hammer, or the hands allows the percussionist to excite the instrument with varying forces and velocities, and to alter the timbre of the produced sound by changing the shape, hardness, weight, and stiffness of the tool. All these features make drumming gestures a very rich subject of study which has not been previously explored before in computer animation, both for the expressiveness and the variability of the performance, and for the control of synthesized sounds.

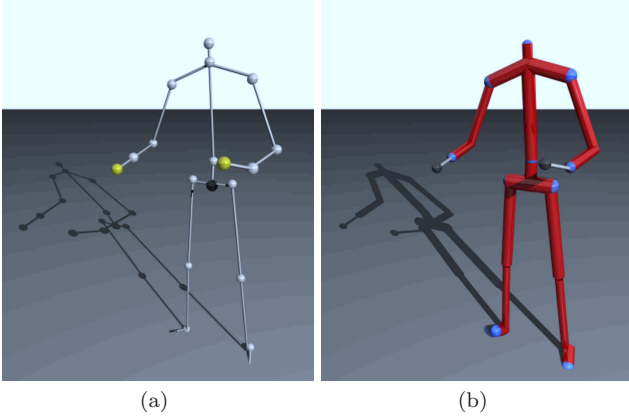


Fig. 3 Virtual character model: merging of kinematic and dynamic representations. (a) The underlying kinematic representation is represented by the implicit limbs (white lines) linked by a set of joints (white spheres), among which arm end-effectors (yellow spheres) are of special interest in *HIMC* as they represent the extremities of the drumsticks. (b) The underlying dynamic representation is characterized by explicit limbs (red capsule-shaped rigid bodies) linked by mechanical joints (blue spheres).

We chose to work on timpani gestures which have the specificity to involve the whole upper-body during performance, and are characterized by movements of large amplitude during the preparatory phase. For this class of gestures we have captured a large amount of data expressing different styles and nuances. Our captured database covers forty five sequences of drumming, for three percussionists, with five playing modes (*legato*, *tenuto*, *accent*, *vertical accent* and *staccato*), and three impact locations on the timpani, each sequence containing at least 10 strikes, both for individual training exercises and improvisation sequences. This study required the participation of two students and one professional percussionists to make a selective choice of drumming sequences and to record the different performances.

3 Motion simulation: virtual character modeling and motion control

We describe in this section the approach that we adopt for physically modeling and controlling percussion motion from pre-recorded motion capture data. It involves a muscular-skeleton model of the virtual character depicted in figure 3, by merging the two traditional kinematics and dynamics representations [35]. Our novel motion control paradigm takes advantage of this double representation by the definition of a hybrid controller driven by motion capture data according to two control modes. The first control scheme uses simple *ID* controllers similar to previous contributions, whereas the second *HIMC* mode involves a novel hybrid motion

Tab. 1 Muscular-skeleton model: body mass percentages, joint ellipsoid parameters.

Body	m [%]	Joint	$\mathbf{l} [\lambda, (e_x, e_y, e_z)]$
Head	4	Neck	5.7, (0.111, 0.993, 0.045)
Neck	4		2.9, (-0.978, 0.117, -0.173)
Trunk	46		2.3, (0.177, 0.025, -0.984)
Up. arms	4	Shoulder	4.4, (0.291, 0.018, -0.957)
Forearms	2		2.7, (-0.955, -0.062, -0.291)
Hands	1		0.1, (0.064, -0.998, 0.001)
Sticks	0.5	Elbow	14.6, (0.337, -0.684, 0.647)
Pelvis	18		8.9, (-0.889, -0.004, 0.458)
Up. legs	14		4.4, (-0.311, -0.729, -0.609)
Lo. legs	5	Wrist	19.9, (0.231, -0.932, 0.278)
Feet	1.5		9.1, (0.003, -0.285, -0.958)
			2.1, (-0.973, -0.222, 0.063)

control scheme including a cascaded process of *IK* and *ID* controllers. This latter allows for the control of the physical model of the virtual character solely by the specification and editing of drumstick extremity trajectories.

3.1 Virtual character modeling

The muscular-skeleton model of the virtual performer is composed of two skeleton layers: a dynamic layer and a kinematic layer, equation (1).

$$\begin{aligned} \mathbf{D} &= [\mathbf{s}^{\mathbf{D}} = \{m, s, d, \mathbf{I}\}_{[1,N]}, \mathbf{j}^{\mathbf{D}} = \{\mathbf{k}_s, \mathbf{k}_d\}_{[1,M]}] \\ \mathbf{K} &= [\mathbf{j}^{\mathbf{K}} = \{\mathbf{q}, \dot{\mathbf{q}}\}_{[1,M]}, \mathbf{l}^{\mathbf{K}} = \{\lambda, \mathbf{e}\}_{[1,M]}] \end{aligned} \quad (1)$$

The dynamical skeleton \mathbf{D} models the physical properties of the virtual character, and is composed of a set of rigid bodies $\mathbf{s}^{\mathbf{D}}$ articulated by mechanical joints $\mathbf{j}^{\mathbf{D}}$, figure 3(b) shows the considered capsule-shaped skeleton. Motion capture data of percussion performances are used to physically model and parameterize both the anthropometry and the mechanical joints of the virtual character, making a direct yet simplified correspondence between the recorded performer and the virtual character.

The physical properties of each rigid body composing the virtual character, such as mass (m), size of the limbs (s), density (d) and inertia matrix (\mathbf{I}) are consistent with the anthropometry extracted from motion capture data. Table 1 specifies the mass percentage of each rigid body as regards to the total mass of the virtual character, from which the other dynamic parameters are derived using human anthropometry knowledge, for instance [15]. Note that drumsticks are integral parts of the dynamic model \mathbf{D} of the virtual character, so that they can be considered as proper body limbs,

Alg. 1 Quaternionic joint limits.

In	\mathbf{Q}	Joint quaternionic trajectory
Out	(λ, \mathbf{e}) 1	Eigen values and axes Joint limits (bounding ellipsoid)
— Mean of \mathbf{Q} $\bar{\mathbf{q}} \leftarrow \text{mean}(\mathbf{Q})$ — Motion transposition in the tangent space of $\bar{\mathbf{q}}$ $\mathbf{Q}^* \leftarrow \log(\bar{\mathbf{q}} * \mathbf{Q})$ — Eigen values and axes $(\lambda, \mathbf{e}) \leftarrow \text{SVD}(\mathbf{Q}^*)$ — Joint limits: bounding ellipsoid $1 \leftarrow \text{ellipsoid}(\lambda, \mathbf{e})$		

consequently upper-body end-effectors will be referred to as drumsticks extremities.

Mechanical joints are parameterized by proportional (\mathbf{k}_s) and derivative (\mathbf{k}_d) coefficients that traditionally represent the tension/relaxation between rigid bodies in PD inverse dynamics control.

The kinematic layer \mathbf{K} of the virtual character as shown in figure 3(a) represents the observation of the kinematic state of \mathbf{D} , *i.e.* the linear-angular positions and velocities $(\mathbf{q}, \dot{\mathbf{q}})$ of the joints articulating the dynamic representation. State features \mathbf{q} and $\dot{\mathbf{q}}$ are obtained by integrating the set of motion equations over time through a forward dynamic scheme, using the Lagrangian formalism $\frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}) - \frac{\partial L}{\partial q_i} = Q_i$, q_i and \dot{q}_i being the generalized coordinates (quaternion angular position and angular velocity), and Q_i the generalized effort for each degree of freedom i of the system.

In addition, each mechanical joint has three rotational degrees of freedom, restricted to the angular limits of the human body, in order to avoid non realistic motion. Two formulations are available to extract the "lower" and "upper" limits of the angular joints, based on a statistical analysis of the motion capture data. The first formulation uses the Euler angle representation, and computes basic statistical features to characterize joint limits. The singularity of this representation is however frequent (*gimbal lock*), therefore we propose another formulation based on the quaternionic representation presented in [20].

Joint limits are processed in the quaternion space, as detailed by algorithm 1. The rotational (quaternionic) trajectory \mathbf{Q} of a joint over time is transposed in the tangential space of its quaternionic mean $\bar{\mathbf{q}}$, in which a Singular Value Decomposition is applied. The resulting eigen values and axes (λ, \mathbf{e}) are used for representing the joint limits 1 of a joint by the specification of the

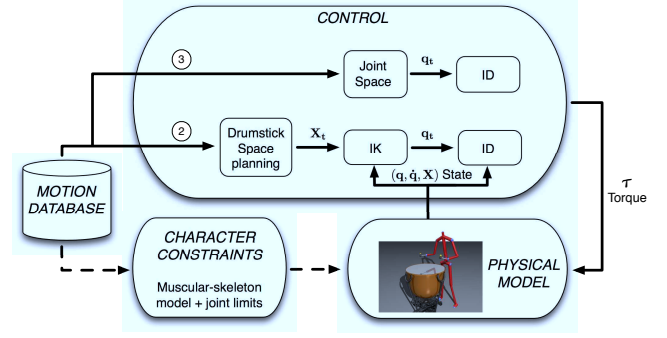


Fig. 4 Hybrid physics-based motion control and synthesis. The *Motion Database* is used for the physics modeling of the virtual character, as well as for expressing two modes of *Control*. The two modes correspond to pipelines (2) and (3) depicted in figure 1. These allow for the physics-based motion tracking, (3) in the *Joint Space* (*ID*) from angular quaternionic trajectories \mathbf{q}_t , or (2) in the *Drumstick Space* (*HIMC*) from end-effector trajectories \mathbf{X}_t . *HIMC* involves the combination of inverse kinematic (*IK*) and inverse dynamic controllers (*ID*).

bounding ellipsoid around the considered quaternionic mean.

During the simulation, the respect of joint limits from bounding ellipsoids is then achieved by verifying if each joint orientation transposed in its respective mean tangent space is valid (*i.e.* contained in ellipsoid). If it is not valid, torques are computed so that mechanical joints achieve the orientation corresponding to the tangent space vector optimizing the quadratic distance on the considered ellipsoid surface. This minimum-distance in the tangent space corresponds to the minimum-path on the unit sphere in the quaternion space. Table 1 also specifies the eigen values and axes representing the bounding ellipsoid for the upper-body joints, *resp.* neck, shoulder, elbow and wrist. These parameters have been obtained from the different percussion recordings, allowing to constrain these parameters for the specific task of percussion.

3.2 Motion control

We propose two ways for realizing the motion control of the muscular-skeleton model, as illustrated in figure 4. The first control paradigm proceeds to the traditional tracking of motion capture data in the joint space (quaternionic trajectories). The novel *HIMC* scheme that we propose is motivated by the expressive factor of drumstick control, and proceeds to the tracking of end-effector trajectories in the 3D Cartesian space. While tracking motion capture data in the joint space requires *ID* control, *HIMC* tracking in the end-effector space requires a cascaded collaboration of both *IK* and *ID*

controllers. In the latter case, both inversion processes are strongly linked.

3.2.1 ID motion control

ID control is related to tracking motion capture data by using PD controllers. Kinematic postures of the joints composing the virtual character (quaternionic trajectories \mathbf{q}_t) are first extracted from motion capture data, and used to drive the motion of the physics-based model of the virtual character.

We use traditional PD feedback controllers, modeled as damped springs and parameterized by manually-tuned stiffness and damping coefficients ($\mathbf{k}_s, \mathbf{k}_d$). Knowing the current state of the mechanical joints ($\mathbf{q}, \dot{\mathbf{q}}$) and the angular target \mathbf{q}_t to be reached, the needed torque $\boldsymbol{\tau}$ is computed and exerted on the corresponding articulated bodies according to equation (2).

$$\boldsymbol{\tau} = \mathbf{k}_s \cdot (\mathbf{q}_t - \mathbf{q}) - \mathbf{k}_d \cdot \dot{\mathbf{q}} \quad (2)$$

3.2.2 HIMC: Hybrid Inverse Motion Control

The Hybrid Inverse Motion Control (HIMC) is a more intuitive physics control of the virtual character combining IK and ID controllers. Instead of directly tracking quaternionic trajectories (\mathbf{q}_t) from the motion capture database, HIMC consists in tracking end-effector positions in the 3D Cartesian space (drumsticks extremities \mathbf{X}_t for arms, for instance).

Knowing the Jacobian \mathbf{J}_q of the system to be controlled, as well as its Cartesian target \mathbf{X}_t and current state \mathbf{X} , an IK method computes a kinematic posture \mathbf{q}_t according to equation (3). The kinematic posture \mathbf{q}_t is then used similarly to the previous point as the desired input of PD controllers, thus providing the required torques $\boldsymbol{\tau}$ to control the physical character, as shown in figure 4.

$$\begin{aligned} \Delta \mathbf{q}_t &= \mathbf{J}_q^{-1} \cdot (\mathbf{X}_t - \mathbf{X}), \quad \mathbf{q}_t = \mathbf{q} + \Delta \mathbf{q}_t \\ \boldsymbol{\tau} &= \mathbf{k}_s \cdot (\mathbf{q}_t - \mathbf{q}) - \mathbf{k}_d \cdot \dot{\mathbf{q}} \end{aligned} \quad (3)$$

Algorithm 2 shows the principle of the HIMC control scheme, composed of two nested loops. The first one concerns the task specification (drumstick extremity position). Between two task updates, the second loop involves the cooperation of the IK and ID formulations. For example, for a task frequency of 250 Hz and a simulation frequency of 10000 Hz, the second simulation loop is called $10000/250=40$ times, which is experimentally sufficient for the IK convergence, and consequently the ID convergence.

Alg. 2 HIMC combining IK and ID controllers.

	F_T, F_S	Task and simulation frequencies
In	$(\mathbf{q}, \dot{\mathbf{q}})$	Current kinematic posture
	\mathbf{X}	Current end-effector position
	\mathbf{X}_t	Targetted end-effector position
Out	\mathbf{J}_q	Jacobian matrix
	\mathbf{q}_t	Targetted kinematic posture
	$\boldsymbol{\tau}$	Resulting torques

```

while Read task at  $F_T$  do
     $\mathbf{X}_t \leftarrow$  New drumstick extremity
     $steps \leftarrow 0$ 
    repeat
        — Inverse Kinematics
         $\mathbf{J}_q \leftarrow$  ComputeJacobian( $\mathbf{q}$ )
         $\mathbf{q}_t \leftarrow$  InverseKinematics( $\mathbf{J}_q, \mathbf{X}, \mathbf{X}_t$ )

        — Inverse Dynamics
         $\boldsymbol{\tau} \leftarrow$  InverseDynamics( $\mathbf{q}, \dot{\mathbf{q}}, \mathbf{q}_t$ )

        — Forward Dynamics
        ApplyTorques( $\boldsymbol{\tau}$ )
        Simulation()
         $steps \leftarrow steps + 1$ 
    until  $steps = F_S/F_T$ 
end

```

The modularity of the presented approach for combining IK and ID controllers allows the consideration of any method, so that our system supports at the moment several IK formulations that can be combined with simple ID formulations: transpose, pseudo-inverse, damped-least-squares. Recent advances using learning techniques [18,2] could also be considered.

Using an IK formulation necessitates the computation of the Jacobian matrix, underlining the need of the double kinematic and dynamic representation of the virtual character. The main difficulty with the coupling of both kinematic and dynamic controllers is that the convergence of the IK algorithm is added to the difficulty of tuning the parameters of the PD dynamic controllers. However, once adequate PD parameters are found, this approach enables the manipulation of motion capture data in the 3D Cartesian space (configuration \mathbf{X}_t) instead of the joint space (\mathbf{q}_t), which is more consistent and intuitive for controlling percussion motion. Using end-effector trajectories such as drumstick extremities is a means of compressing the motion, and may therefore facilitate the edition of percussion gestures to be synthesized.

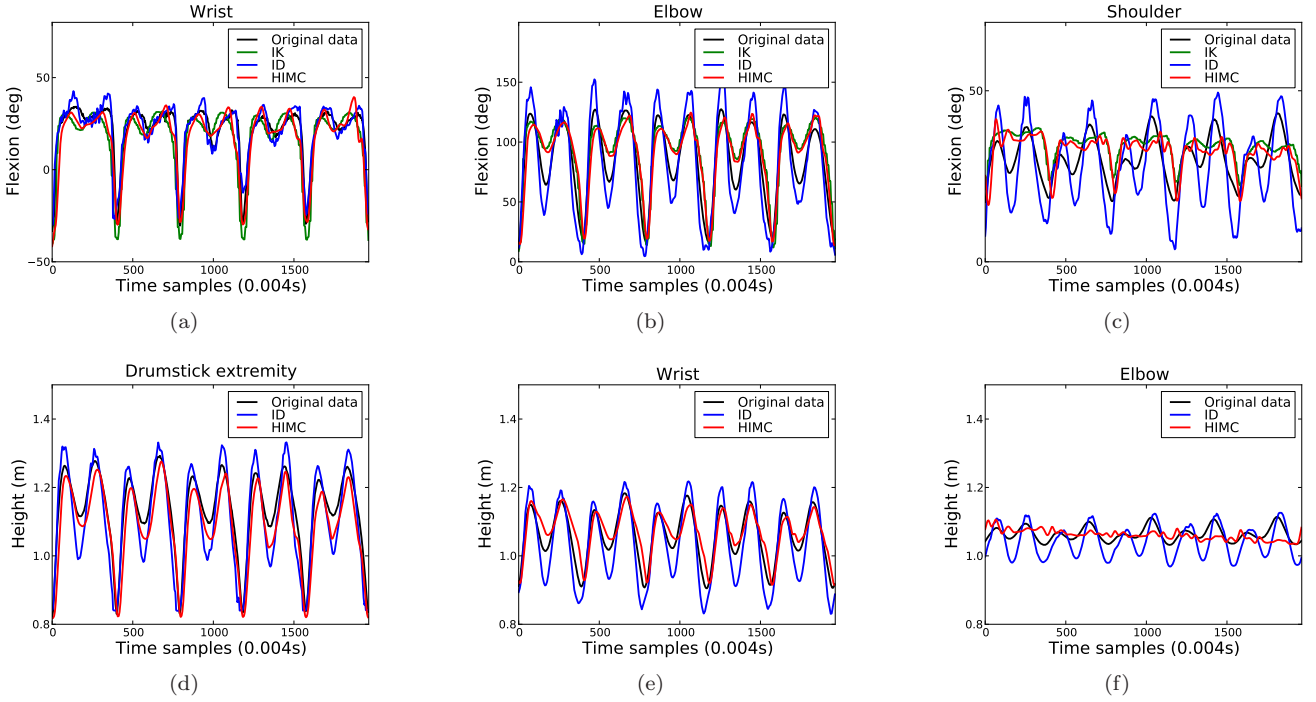


Fig. 5 Comparison of captured (black) and simulated trajectories for several percussion strokes, using *IK* (green), *ID* (blue) and *HIMC* (red) – (a-b-c) wrist, elbow and shoulder flexion angles, (d-e-f) drumstick extremity, wrist and elbow height positions.

3.3 Results

The physical model of the virtual character is composed of 19 joints totalling 57 degrees of freedom. The presented results focus more specifically on the upper-body, which counts 11 bodies articulated by 11 joints totalling 33 degrees of freedom. The *Open Dynamics Engine* (*ODE*) physics simulation toolkit¹ is used for the forward dynamic simulation of the motion equations. The *ID* implementation used in these experiments also exploits *ODE*'s internal force/torque maxima, for preserving simulation stability while not sacrificing simulation quality with over-damped behaviours. The *HIMC* tracking mode can take advantage and can be combined to various inverse kinematic methods. These results have been obtained using an implementation of the damped-least-squares method [37], a simple yet more robust adaptation of the pseudo-inverse method regarding the singularity of the inverse kinematic problem.

The results presented in this section compare the *HIMC* control mode to both traditional *IK* and *ID* control modes. It should be noted that to make such a comparison possible between *HIMC* and *ID* control modes, parameters that could drastically change the resulting simulations have been kept constant. Such parameters include for instance the simulation rate, the recorded

motion to be simulated, as well as the parameterization of the damped springs composing the dynamic layer of the virtual character.

A qualitative evaluation is first conducted. The results obtained by the two tracking modes used to physically control the virtual percussionist are compared: a) motion capture tracking in the joint space, involving only *ID* controllers, and b) motion capture tracking in the Cartesian space, using the *HIMC* combination of *IK* and *ID* controllers. Then we extensively and quantitatively evaluate the two tracking modes by quantifying the errors made during the simulation of various motion variations in each case.

A critical issue to consider is the metric used for evaluating the effectiveness of *HIMC* compared to traditional *ID* control. Previous works have addressed the question of the user sensitivity to various motion metrics [30,31]. In this work, the ballistic nature of percussion motion does however make it easy to conduct an extensive evaluation by focusing on drumstick extremity trajectories. We also show that our solution leads to a coherent motion in the joint space.

3.3.1 Qualitative evaluation

For this qualitative evaluation, we ran the simulation of a set of pre-recorded percussion motions. Figure 5 compares the original motion to the resulting simulated

¹ *Open Dynamics Engine*, <http://www.ode.org>

motion using the three considered control modes: *IK* only, *ID* only and *HIMC*.

The hybrid control scheme tracks a sequence of percussion gestures for synthesizing whole arm movements only from the specification of the drumstick tips. Figures 5(a), 5(b) and 5(c) present the comparison between original data and data generated by the three control modes for wrist, elbow and shoulder flexion angle trajectories. These results show that data generated by the *IK* involved in the hybrid control mode are consistent with real ones, especially for wrist and elbow trajectories. *IK* techniques do not generally guarantee that angular data are exactly following mocap data, however *IK* gives coherent results thanks to its embedding with *ID* in the core control process, and thanks to a tracking process using close successive targets. Data generated by our hybrid solution is even more accurate compared to *ID* control, the latter shows indeed a motion exaggeration that is avoided by our solution. More specifically for the shoulder flexion angle, it seems that the *IK* process alters the original motion. Joints closer to the skeleton root are generally considered as more important than the others, for instance for locomotion. However this is not necessary the case for the considered percussion tasks. When the percussionist faces his instrument, the position of the root does not move very much, so that the motion of his arms (elbow, wrist) become more important than shoulder motion. This loss is therefore of lower importance for percussion motion, as argued in [6,8] that wrist and elbow angle trajectories are important degrees of freedom in percussion (timpani) arm mechanisms, in comparison to shoulder angle trajectories.

We also present the comparison of the two control modes (*ID* control only and *HIMC*) for height trajectories of the drumstick extremity, wrist and elbow in figures 5(d), 5(e) and 5(f). One interesting issue is the accuracy of the hybrid control mode compared to the simple *ID* control. The motion exaggeration previously observed in joint flexion trajectories in the case of *ID* control is propagated on joint position trajectories. Conversely, *HIMC* leads to a more accurate tracking in joint position trajectories, as shown in the animation (Online Resource 1). This observation lies in the fact that the convergence of motion capture tracking is processed in the joint space in the case of *ID* control, adding and amplifying multiple errors on the different joints and leading to a greater error than processing the convergence in the Cartesian space for the hybrid control.

One limitation however arises in elbow position trajectories. While the *ID* control mode results again in an amplified motion, our hybrid control solution re-

sults in a loss in the motion pattern of original data. The elbow position trajectory resulting from the hybrid control mode leads to a stiffer and somewhat constant motion. This can be related to the inverse kinematic process which does not include any secondary goals, as already pointed out in [21]. This may be explained also by the fact that the *IK* algorithm favors the accuracy of the most distal joints compared to joints at the basis of the articulated chain. Another reason for this limitation could be the influence of the physical simulation on the convergence of the inverse kinematic scheme.

Although such a cascaded chain of *IK* and *ID* controllers runs in real-time, the main drawback of this improvement is however the additional computational cost of the *IK* algorithm which is processed at every simulation step. It nevertheless provides a more flexible motion edition technique for controlling a physics-based virtual character solely by end-effector trajectories. It also provides a consistent control scheme as regards to current knowledge about the control mechanisms occurring in real percussion performances [13,8].

3.3.2 Quantitative evaluation

An extensive evaluation of *ID* and *HIMC* control modes is conducted, taking into account the test of the two control algorithms on 100 simulation trials for various motion variations (from top to bottom: *legato*, *tenuto*, *accent*, *vertical accent* and *staccato*).

Figure 6(a) shows examples of depth/height phase trajectories of the drumstick extremity position for each motion variation. Motion variations present drastically different dynamics in drumstick motion, especially regarding its depth and height amplitudes. For example, *accent* variation differs from the others through a stiffer motion, characterized by a higher height amplitude and a more reduced range of motion in the sagittal direction. Conversely *vertical accent* and *staccato* variations are special compared to the others as they are characterized with a much greater depth motion amplitude. These differences in motion dynamics play an important role in the expressiveness of percussion motion, as illustrated in the corresponding animation (Online Resource 1).

Results of the quantitative evaluation are illustrated in figure 6(b). The experience consisted in the generation of 100 simulation trials for each motion variation. Each simulation trial leads to the computation of the root mean square error (*RMS-E*) and standard deviation (*RMS-SD*) as regards to the drumstick extremity from the corresponding motion initially chosen in the motion capture database. From the results presented in figure 6(b), one can easily conclude that the cascaded

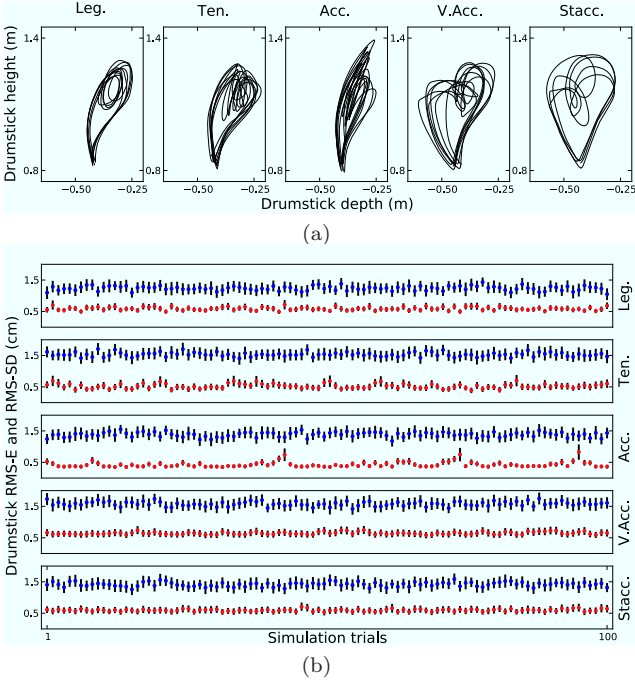


Fig. 6 (a) Depth/height phase trajectories of drumstick extremity for 5 motion variations, from left to right: *legato*, *tenuto*, *accent*, *vertical accent* and *staccato*. (b) Comparison of root mean square error ($RMS-E$) and standard deviation ($RMS-SD$) between captured and simulated drumstick trajectories for various motion variations, using the two motion control schemes: ID and $HIMC$. $RMS-E$ (blue for ID , red for $HIMC$) and $RMS-SD$ (black) are obtained from 100 simulation trials for each motion variation (from top to bottom: *legato*, *tenuto*, *accent*, *vertical accent* and *staccato*).

hybrid combination of IK and ID controllers leads to a more accurate minimization of the error made on drumstick extremity trajectories of about 2% of motion amplitude, compared to simple ID controllers. The standard deviation along simulation trials are also more reduced for the hybrid motion paradigm compared to the ID control. Such a difference of 2% in the average errors between the two control modes could be considered as negligible. However, drumstick extremity information around beat impacts can change drastically the nature of the resulting sounds, that is why such amelioration is of paramount importance for proposing compelling contact information both for motion simulation and sound synthesis.

3.3.3 Discussion

Two interesting points about $HIMC$ that are worth mentioning are the interest of drumstick extremity control for motion simulation on the one hand, and on the other hand for sound synthesis.

$HIMC$ is a means of reducing the dimensionality of motion capture data for controlling a physically respon-

sive virtual character. Only 3D Cartesian trajectories are necessary compared to traditional high-dimensional body postures. $HIMC$ is therefore characterized by a meaningful, intuitive yet real-time scheme for physics tracking. One critical point in $HIMC$ is whether or not to embed IK inside the physics control loop (*cf.* algorithm 2). A first approach could consist in separating the IK process from the physical simulation, resulting in the preprocess of end-effector targets before the simulation. Our choice with $HIMC$ is to embed and combine IK and ID at the physical simulation level. It does not hypothesize any pre-processing while achieving real-time simulation. $HIMC$ makes therefore room for the real-time modification of the desired end-effector tasks performed by the virtual character depending on physical interactions occurring during the simulation, for instance the membrane displacement fed back to the virtual performer.

$HIMC$ is also carefully designed to fulfill some of the requirements of sound synthesis. Its physical nature provides any sound synthesis model with the essential contact parameters that give an adequate sound feedback to motion simulation. Moreover $HIMC$'s capability to control and modify in real-time end-effector tasks is important in the case of percussive sound synthesis. Physics-based sound synthesis of percussive sounds are generally based on simulating vibrations. Particularly among membranophones, membrane vibrations play an important role during the drumstick/membrane interaction and finally in the resulting sound. Vibrations depend on membrane physical properties (mass, tension, Young's modulus), as well as the applied force, and in return vibration modes will act on and disturb differently the drumstick. With $HIMC$, it is therefore possible to explore this phenomenon, for instance by exploring models accounting for the drumstick reaction towards membrane vibrations, although this was not implemented in the current simulation.

4 Interaction and sound synthesis

In this section, we present the explicit interaction of percussion motion simulations with a modal/mechanical model of a drum membrane. We then propose a general architecture for running simultaneously and asynchronously the physical simulation of percussion gestures, graphics and sound rendering processes, as well as handling their interaction in real-time. This architecture is effective for managing heterogeneous data types characterized by various frame rates, and for specifying at the physics level the mapping between these modalities.

Tab. 2 Mapping parameters between $\mathbf{f}_{\text{motion}}$ and $\mathbf{f}_{\text{sound}}$.

$\mathbf{f}_{\text{motion}}$	$\mathbf{f}_{\text{sound}}$
Impact location (x_0, y_0)	Membrane inner properties (M : e.g. size, tension)
Contact shape (ϕ)	
Contact compression (c)	
Force amplitude (F_m)	Force profile (F_s)

4.1 Motion-sound physics mapping

In this paper, we consider the modal synthesis technique, which is an efficient way of representing the vibrations of resonating objects as the motion simulation of systems composed of masses connected with springs and dampers. A practical mapping between the force generated by the motion simulation $\mathbf{f}_{\text{motion}}$ and the interaction force $\mathbf{f}_{\text{sound}}$ acting on a modal/mechanical model of a drum membrane is shown in table 2.

The drum membrane model is governed by the vertical motion $\mathbf{z} = [z_i]_{i \in [1, N]}$ of N mechanical oscillators of mass $\mathbf{M} = [m_i]_{i \in [1, N]}$ and spring coefficients $\mathbf{K} = [k_i]_{i \in [1, N]}$, that can be excited by an external force $\mathbf{f}_{\text{sound}} = [f_i]_{i \in [1, N]}$ according to equation (4). This model is characterized by coupled equations since matrices \mathbf{M} and \mathbf{K} are not necessarily diagonal.

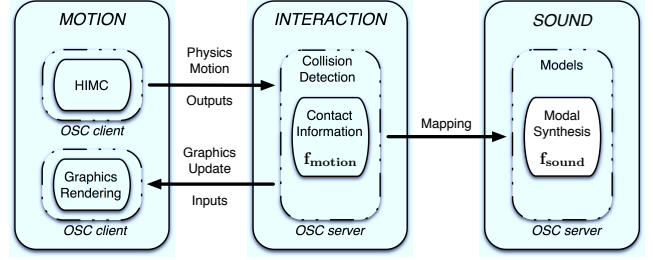
$$\mathbf{M} \cdot \ddot{\mathbf{z}}(t) + \mathbf{K} \cdot \mathbf{z}(t) = \mathbf{f}_{\text{sound}} \quad (4)$$

$$\begin{aligned} \mathbf{M}_{\mathbf{S}} \cdot \ddot{\mathbf{q}}(t) + \mathbf{K}_{\mathbf{S}} \cdot \mathbf{q}(t) &= \mathbf{S}^T \cdot \mathbf{f}_{\text{sound}} \\ \text{with : } \mathbf{S} &= [\mathbf{s}_1, \dots, \mathbf{s}_N], \mathbf{z} = \mathbf{S} \cdot \mathbf{q} \\ \mathbf{M}_{\mathbf{S}} &= \mathbf{S}^T \cdot \mathbf{M} \cdot \mathbf{S} \\ \mathbf{K}_{\mathbf{S}} &= \mathbf{S}^T \cdot \mathbf{K} \cdot \mathbf{S} \end{aligned} \quad (5)$$

$$\begin{aligned} F_s &= k \cdot [c(x_0, y_0, M, t)]^\alpha \\ \text{with : } k &\propto F_m \\ c(x_0, y_0, t) &= p(t) - d(x_0, y_0, M, t) \\ \alpha &= f(\phi) \end{aligned} \quad (6)$$

The factorization of the model governed by equation (4) leads to a solution of the form $\mathbf{z} = [\mathbf{s}_1, \dots, \mathbf{s}_N] \cdot \mathbf{q}$ where $\mathbf{q} = [q_i]_{i \in [1, N]}$ are sinusoidal oscillators (or modes) and $\mathbf{s}_i = [\mathbf{s}_{i,j}]_{j \in [1, N]}$ are referred to as the *modal shapes* of the system. Including the latter expression in equation (4) which leads to considering the modal shapes $[\mathbf{s}_1, \dots, \mathbf{s}_N]$ as an orthogonal basis. The expression of the initial problem yields an uncoupled equivalent formulation in this orthogonal basis, equation (5), where $\mathbf{M}_{\mathbf{S}}$ and $\mathbf{K}_{\mathbf{S}}$ are diagonal matrices. Equation (5) explicitly shows that the *modal shapes* of the system (matrix \mathbf{S}) defines how external force $\mathbf{f}_{\text{sound}}$ acts on the modes.

Such a physically-based approach enables the direct modeling of contact force impacts. This force mainly depends on the state (displacement and velocity) of the colliding modal objects. A collision detection algorithm

**Fig. 7** Asynchronous server-client architecture of our system.

associated with the motion physics simulation retrieves the physical features $\mathbf{f}_{\text{motion}}$ of any contact event produced when the drum is excited during the simulation. It specifically provides information on impact location (x_0, y_0) , contact shape ϕ , contact compression c , as well as force amplitude F_m . Contact compression is specifically expressed as the difference between drumstick and membrane vertical displacement (*resp.* p and d), with d depending on the impact location and membrane properties M . This contact information can be put in correspondence to the previously presented inputs to the modal/mechanical model of the drum membrane according to a direct one-to-one mapping. Equation (6) details a possible mapping following Hertz's contact law. Force profile F_s acting on the drum model depends both on a force stiffness term k and compression c . Force stiffness k is determined from force amplitude F_m during the simulation, whereas c is powered by a term α that characterizes the contact geometry (ϕ) .

4.2 Asynchronous client-server architecture

Multimodal processes such as graphics and sound rendering are fundamentally different, and achieving a real-time interaction between the two can be hazardous. The first difficulty that appears when managing such media is their discrepancy in time constants: graphics rendering is usually admitted effective at about 33Hz, whereas sound rendering needs a much higher sampling rate around 44kHz. Moreover, the graphics rendering is the visible layer of a far more demanding process, that of the physical simulation, requiring the handling of two other different time rates: the original motion capture time rate (250 Hz) and the time step of the physical simulation (10000 Hz). A solution could consist in running synchronously every manager at the fastest time rate (sound), but such an approach would fall short in real-time considerations.

We propose an asynchronous client-server scheme for handling the interaction between the physics, graphics and sound managers as depicted in figure 7. We adopt the *Open Sound Control* (OSC) communication

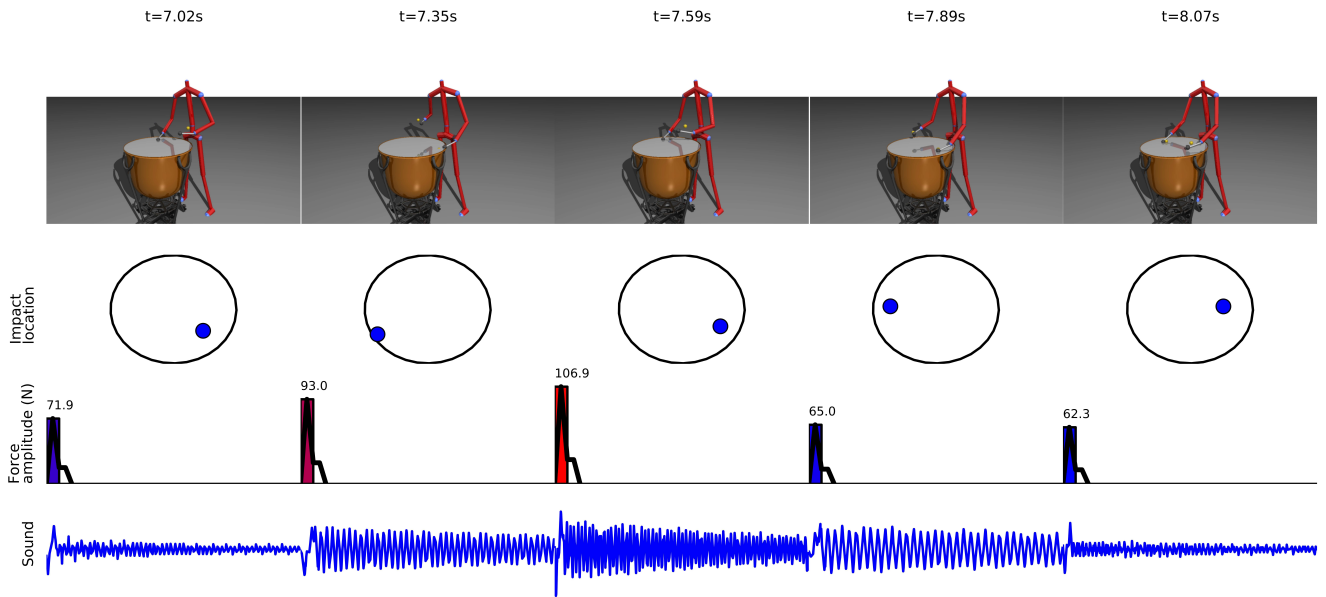


Fig. 8 Application: simulation of timpani motion using the *HIMC* control mode and its interaction with the modal sound synthesis model. The virtual character is controlled by the specification of drumstick extremity trajectories (yellow spheres). An excerpt of a timpani improvisation including various motion and rythm variations is presented, consisting of five percussion strokes with corresponding impact locations on the drum membrane as well as impact forces and resulting sound waveforms.

protocol² which is traditionally used for allowing the real-time exchange of multimedia data flows. The physical simulation and graphics rendering are embedded in *OSC* clients that are registered to the interaction and sound synthesis managers defining *OSC* servers. Such architecture allows for the asynchronous interaction of motion simulation, visual and sound renderings, namely by sending contact information to sound synthesis processes when needed. This also enables the possible distribution of the different managers on distinct platforms, thus reducing the computational cost of each manager and its impact on the others. The asynchronous exchange of data is materialized here by the mapping between the output of motion synthesis and the input of sound synthesis. Events produced during the physical simulation are detected by the interaction manager (collision detection). It is then in charge of processing and propagating these information through the *OSC* protocol so that the graphics and sound managers can react and produce the corresponding sound and visual outputs.

4.3 Application: simulation of a timpani improvisation

An application of our motion control simulation and interaction with modal synthesis is presented in this section, with a special focus on timpani percussion motion (figure 8). An improvisation made of various motion

variations (*legato*, *tenuto*, *accent*, *vertical accent* and *staccato*) and rythmn variations of timpani motion is simulated and shown in the corresponding animation (Online Resource 1).

The 3D rendering of the timpani percussion instrument is the visible representation of the modal/mechanical model of the drum membrane, with which the virtual percussionist can interact. The drum model is characterized by a membrane diameter of 23" and has been modeled using *Modalys*³. The motion control and simulation using *ODE* includes a collision detection module, which once triggered by a collision between drumsticks and the drum membrane sends contact information (impact location, contact shape and compression, force amplitude) both to graphics and sound managers. In this application, motion/interaction and sound managers have been deployed on two different machines following our system architecture, *cf.* figure 7. This results in the real-time visualization of simulated beat impacts on the drum membrane, as well as corresponding sound feedback.

Figure 8 shows an excerpt of percussion impacts, with their corresponding impact locations on the drum membrane as well as impact forces and resulting sound waveforms. Motion control is performed using *HIMC* and implies only the specification of drumstick extremity trajectories (yellow spheres seen in screenshots). The specification of drumstick extremity trajectories facili-

² *Open Sound Control*, <http://www.opensoundcontrol.org>

³ *Modalys*, <http://support.ircam.fr/doc-modalys>

tates the edition of the motion to be simulated, since the control of whole upper-body motion is reduced to two 3D position trajectories. Moreover, *HIMC* control mode from drumstick trajectories allows for the specification of various impact locations on the drum membrane, and therefore produces adequate variations of sound rendering. Figure 8 also shows the capability of our system to take into account varying impact forces. The first three percussion impacts show an increase in force impacts with an adequate increase in sound amplitude. Specifically, on the third impact a force of 106.9N is generated, and as a result a much greater waveform amplitude is synthesized compared to the other impacts. The next impact is characterized by a lower force amplitude of 65N, but takes advantage of the resonance of drum vibrations created by the previous impact to produce a medium sound amplitude, compared to the fifth impact with a similar force (62.3N), but with a lower sound amplitude.

5 Conclusion and Future Works

In this paper, we proposed a physically-enabled environment in which a virtual percussionist can be physically controlled and interact in real-time with physics-based sound synthesis processes. The control of the virtual character uses motion data recorded during real percussion performances, therefore guaranteeing to maintain the main characteristics of human motion data while keeping the physical coherence of the interaction with the simulated instrument. For the range of our timpani expressive gestures, we have shown that the motion control method is efficient and robust. This can be explained by the incorporation within our control scheme of constraints statistically computed from motion captured data, that is biomechanical parameters and joint limits. *HIMC* control mode, combining *IK* and *ID* controllers, leads to a more accurate and controllable simulation as regards to drumstick extremity trajectories. The generation of consistent poses over time is also ensured by the dynamic tracking of these trajectories. Finally, driving the virtual character using such data as input provides users with a more intuitive way of composing the motion to be simulated, by editing and manipulating these trajectories as a gesture partition. We also proposed a modal/mechanical model of a drum membrane with which the physics-based model of the virtual percussionist can interact. Finally, our asynchronous client-server architecture takes advantage of motion and sound physical formulations, generating in real-time virtual percussion performances.

This framework has been implemented in a modular way, so that we can substitute each module dedicated

to motion or sound synthesis. In the near future, other more elaborate schemes of inverse kinematics will be used, in particular those exploiting learning of motion captured data. A major work will also involve the extension of *HIMC* to full-body motion synthesis to the light of recent achievements. Balance strategies similar to [24] could therefore be used for simulating our virtual performer playing with a set of several percussion instruments, while keeping the effectiveness of *HIMC* for simulating upper-body motion. This would also allow to explore a large variety of percussion gestures and also different styles of percussionists. Among other perspectives of such work is the improvement of the interaction between motion simulation and sound synthesis. This can involve the design of a more powerful collision detection module. If we consider the physical simulation of the instrument, it might be interesting to include mechanical interactions, such as hammered or plucked interactions, or even to model the mechanical structure of a finger or a hand. We also plan to extend the mapping possibilities between motion parameters and sound input parameters, and to develop a user interface for editing percussion motion in the 3D Cartesian space. This would lead to the possibility of building a new pedagogical tool dedicated to percussion performance analysis and simulation.

6 Acknowledgments

This work was partially supported by the *Natural Sciences and Engineering Research Council of Canada (Discovery and Special Research Opportunity grants)*, the *Agence Nationale de la Recherche* and the *Pôle de Compétitivité Images & Réseaux* (France).

References

1. B. Allen, D. Chu, A. Shapiro, and P. Faloutsos. On the Beat! Timing and Tension for Dynamic Characters. In *Proc. of the SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 239–247, 2007.
2. M. Aubry, F. Julliard, and S. Gibet. Modeling Joint Synergies to Synthesize Realistic Movements. In *Gesture in Embodied Communication and Human Computer Interaction*, LNCS 5934, pages 231–242. Springer Verlag, 2010.
3. N. Bonneel, G. Drettakis, N. Tsingos, I. Viaud-Delmon, and D. James. Fast Modal Sounds with Scalable Frequency-Domain Synthesis. *Transactions on Graphics*, 27(3):1–9, 2008.
4. A. Bouénard, S. Gibet, and M. M. Wanderley. Enhancing the Visualization of Percussion Gestures by Virtual Character Animation. In *Proc. of the International Conference on New Interfaces for Musical Expression*, pages 38–43, 2008.
5. A. Bouénard, S. Gibet, and M. M. Wanderley. Hybrid Motion Control combining Inverse Kinematics and Inverse Dynamics Controllers for Simulating Percussion Gestures. In *Proc. of*

- the International Conference on Computer Animation and Social Agents*, pages 17–20, 2009.
6. A. Bou  nard, M. M. Wanderley, and S. Gibet. Analysis of Percussion Grip for Physically-based Character Animation. In *Proc. of the International Conference on Enactive Interfaces*, pages 22–27, 2008.
 7. A. Bou  nard, M. M. Wanderley, and S. Gibet. Advantages and Limitations of Simulating Percussion Gestures for Sound Synthesis. In *Proc. of the International Computer Music Conference*, pages 255–261, 2009.
 8. A. Bou  nard, M. M. Wanderley, and S. Gibet. Gesture Control of Sound Synthesis: Analysis and Classification of Percussion Gestures. *Acta Acustica united with Acustica*, 96(4):668–677, 2010.
 9. A. Bou  nard, M. M. Wanderley, S. Gibet, and F. Marandola. Virtual Gesture Control and Synthesis of Music Performances: Qualitative Evaluation of Synthesized Timpani Exercises. *Computer Music Journal*, 35(3), 2011. In press.
 10. M. Cardle, S. Brooks, Z. Bar-Joseph, and P. Robinson. Sound-by-Numbers: Motion-Driven Sound Synthesis. In *Proc. of the SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 349–356, 2003.
 11. R. Corbett, K. van den Doel, J. Lloyd, and W. Heidrich. Timbrefields: 3d Interactive Sound Models for Real-Time Audio. *Presence: Teleoperators and Virtual Environments*, 16(6):643–654, 2007.
 12. S. Coros, P. Beaudoin, and M. van de Panne. Robust Task-based Control Policies for Physics-based Characters. *Transactions on Graphics*, 28(5):1–9, 2009.
 13. S. Dahl. Playing the Accent: Comparing Striking Velocity and Timing in Ostinato Rhythm Performed by Four Drummers. *Acta Acustica united with Acustica*, 90(4):762–776, 2004.
 14. M. DaSilva, F. Durand, and J. Popovi  . Linear Bellman Combination for Control of Character Animation. *Transactions on Graphics*, 28(3):1–10, 2009.
 15. W. Dempster and G. Gaughran. Properties of Body Segments based on Size and Weight. *American Journal of Anatomy*, 120(1):33–54, 1967.
 16. P. Faloutsos, M. van de Panne, and D. Terzopoulos. Composable Controllers for Physics-Based Character Animation. In *Proc. of the Annual conference on Computer Graphics and Interactive Techniques*, pages 251–260, 2001.
 17. W. Gaver. How Do We Hear in the World? Explorations in Ecological Acoustics. *Ecological Psychology*, 5(4):285–313, 1993.
 18. S. Gibet and P. Marteau. Expressive Gesture Animation Based on Non Parametric Learning of Sensory-Motor Models. In *Proc. of the International Conference on Computer Animation and Social Agents*, pages 465–472, 2003.
 19. J. Hodgins, W. Wooten, D. Brogan, and J. O’Brien. Animating Human Athletics. In *Proc. of the Annual conference on Computer Graphics and Interactive Techniques*, pages 71–78, 1995.
 20. M. Johnson. *Exploiting Quaternions to Support Expressive Interactive Character Motion*. PhD thesis, Massachusetts Institute of Technology, USA, 2003.
 21. C. Klein and C.-H. Huang. Review of Pseudoinverse Control for use with Kinetically Redundant Manipulators. *Transactions on Systems, Man and Cybernetics*, 13(3):245–250, 1983.
 22. H. Ko and N. Badler. Animating Human Locomotion with Inverse Dynamics. *IEEE Computer Graphics and Applications*, 16(2):50–59, 1996.
 23. G. Loy. *Musimathics, The Mathematical Foundations of Music*, volume 2. MIT Press, 2007.
 24. A. Macchietto, V. Zordan, and C. Shelton. Momentum Control for Balance. *Transactions on Graphics*, 28(3):80:1–80:8, 2009.
 25. I. Mordatch, M. de Lasa, and A. Hertzmann. Robust Physics-based Locomotion using Low-dimensional Planning. *Transactions on Graphics*, 29(4):71:1–71:8, 2010.
 26. J. O’Brien, C. Shen, and C. Gatchalian. Synthesizing Sounds from Rigid Body Simulations. In *Proc. of the SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 175–181, 2002.
 27. C. Picard, F. Faure, G. Drettakis, and P. Kry. A Robust and Multi-scale Modal Analysis for Sound Synthesis. In *Proc. of the International Conference on Digital Audio Effects*, pages 221–227, 2009.
 28. N. Raghuvanshi and M. C. Lin. Interactive Sound Synthesis for Large Scale Environments. In *Proc. of the International Symposium on Interactive 3D Graphics and Games*, pages 101–108, 2006.
 29. M. Raibert. *Legged Robots that Balance*. MIT Press, 1986.
 30. P. Reitsma and N. Pollard. Perceptual Metrics for Character Animation: Sensitivity to Errors in Ballistic Motion. *Transactions on Graphics*, 22(3):537–542, 2003.
 31. L. Ren, P. Alton, A. Efros, J. Hodgins, and J. Rehg. A Data-driven Approach to Quantifying Natural Human Motion. *Transactions on Graphics*, 24(3):1090–1097, 2005.
 32. A. Shapiro, F. Pighin, and P. Faloutsos. Hybrid Control For Interactive Character Animation. In *Proc. of the International Conference on Pacific Graphics*, pages 455–461, 2003.
 33. T. Takala and J. Hahn. Sound Rendering. In *Proc. of the Annual Conference on Computer graphics and Interactive Techniques*, pages 211–220, 1992.
 34. K. van den Doel, P. Kry, and D. Pai. FoleyAutomatic: Physically-Based Sound Effects for Interactive Simulation and Animation. In *Proc. of the Annual conference on Computer Graphics and Interactive Techniques*, pages 537–544, 2001.
 35. H. van Welbergen, B. van Basten, A. Egges, Z. Ruttkay, and M. Overmars. Real-time Animation of Virtual Humans: a Trade-off between Naturalness and Control. In *Proc. of the Eurographics Workshop on Animation and Simulation*, pages 45–72, 2009.
 36. A. Wagner. Analysis of Drumbeats: Interaction between Drummer, Drumstick and Instrument. Master’s thesis, KTH Royal Institute of Technology, Sweden, 2006.
 37. C. Wampler. Manipulator Inverse Kinematic Solutions based on Vector Formulations and Damped Least Squares. *Transactions on Systems, Man and Cybernetics*, 16(1):93–101, 1986.
 38. P. F. Yang, J. Laszlo, and K. Singh. Layered Dynamic Control for Interactive Character Swimming. In *Proc. of the SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 39–47, 2004.
 39. V. Zordan and J. Hodgins. Tracking and modifying upper-body human motion data with dynamic simulation. In *Proc. of the Eurographics Workshop on Animation and Simulation*, pages 13–22, 1999.
 40. V. Zordan, A. Majkowska, B. Chiu, and M. Fast. Dynamic Response for Motion Capture Animation. *Transactions on Graphics*, 24(3):697–701, 2005.